

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 27, 2001, 16:53:25 ; Search time 4492.43 Seconds
(without alignments)
6649.689 Million cell updates/sec

Title: US-09-830-647-3

Perfect score: 2780
1 aattggcgcgcgcgcctctctg.....aaaaaaaaaactcgcag 2780

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: em_estfun:*

2: em_esthum:*

3: em_estlin:*

4: em_estom:*

5: em_estpl:*

6: em_estba:*

7: em_estro:*

8: em_estov:*

9: em_hlc:*

10: qb_estl:*

11: qb_est2:*

12: qb_hlc:*

13: qb_gss:*

14: em_gss_fun:*

15: em_gss_hum:*

16: em_gss_inv:*

17: em_gss_pln:*

18: em_gss_pro:*

19: em_gss_med:*

20: em_gss_vrt:*

21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	949.2	34.1	950	AL560180	AL560180 AL560180
2	817	29.4	845	AL581669	AL581669 AL581669
3	769	27.7	769	AL548911	AL548911 AL548911
4	768.2	27.6	784	AL573294	AL573294 AL573294
5	756	27.2	1019	AL572094	AL572094 AL572094
6	723.4	26.0	725	AL559244	AL559244 AL559244
7	721.6	26.0	903	AL128881	AL128881 AL128881
8	719	25.9	742	AL580899	AL580899 AL580899
9	713.4	25.7	2338	AK014480	AK014480 Mus muscu
10	712	25.6	712	BF793466	BF793466 602254949
11	710.2	25.5	821	BF496289	BF496289 602538259
12	710	25.5	744	AL948485	AL948485 wq06d08.x

13	692.4	24.9	703	11	BG716928	BG716928 602689445
14	691.6	24.9	767	11	BG717397	BG717397 602689740
15	671.4	24.2	701	10	AM960753	AM960753 EST372824
16	662.8	23.8	950	11	BE888697	BE888697 601513095
17	661.4	23.8	743	10	AI761101	AI761101 w16903.x
18	647	23.3	730	10	AM105287	AM105287 x459f02.x
19	645.4	23.2	659	10	BE613476	BE613476 601504243
20	639.4	23.0	927	11	BG025871	BG025871 602292007
21	627.4	22.6	629	11	BE888268	BE888268 601511791
22	597.8	21.5	603	10	AI824033	AI824033 w12909.x
23	593	21.3	616	11	BE884320	BE884320 601505790
24	586	21.1	701	10	BG722508	BG722508 602693762
25	585.8	21.1	610	10	AI805862	AI805862 t652d11.x
26	585	21.0	585	11	BG724408	BG724408 602689338
27	574.8	20.7	581	11	BG719249	BG719249 602690314
28	572	20.6	670	10	AV645569	AV645569 AV645569
29	568.2	20.4	706	11	BI094492	BI094492 602850523
30	561.6	20.2	839	10	BE564084	BE564084 601348046
31	561	20.2	561	10	AA451820	AA451820 z46f01.x
32	560.4	20.2	578	11	BG724066	BG724066 602697046
33	553	19.9	639	10	AV728361	AV728361 AV728361
34	548.6	19.7	935	11	EG341292	EG341292 602463961
35	548	19.7	549	10	AI923577	AI923577 w175c03.x
36	545.2	19.6	571	10	AU150504	AU150504 AU150504
37	540.2	19.4	760	11	BE219022	BE219022 601882245
38	523.8	18.8	839	10	BE782351	BE782351 601469844
39	517.8	18.6	533	10	AM631180	AM631180 h161905.x
40	493.4	17.7	570	10	AU143969	AU143969 AU143969
41	490.4	17.6	493	10	AI452459	AI452459 t162b12.x
42	488.6	17.6	626	10	BE245831	BE245831 TCBAPIE31
43	484.2	17.4	489	10	AM967974	AM967974 EST380049
44	479	17.2	668	11	BG168560	BG168560 602345241
45	478.4	17.2	740	11	BG965702	BG965702 602830553

ALIGNMENTS

RESULT 1
AL560180 950 bp mRNA
LOCUS AL560180 LTI_FL011_B01 Homo sapiens cDNA clone CS0DG002YE20 5 prime
DEFINITION / mRNA sequence.
ACCESSION AL560180
VERSION AL560180.1 GI:12906394
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source

1..950
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DG002YE20"
/clone_id="LTI_FL011_B01"
/sex="male"
/tissue="B cells from Burkitt lymphoma"
/lab_host="DH10B"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen

Fri Dec 28 08:22:46 2001

us-09-830-647-3.rst

Page 2

```

9600 Medical Center Drive Rockville, Maryland 20850, USA
Fax: (71) 301 610 8371 Email: fielangel@elitech.com URL :
http://fulllength.invitrogen.com
BASE COUNT      303 a      204 c      209 g      232 t      2 others
ORIGIN

```

Query Match	34.1%	Score 949.2	LD 10:	Length 950:
Best Local Similarity	99.8%	Pred. No. 3.9e-144:		
Matches 948:	Conservative 2:	Mismatches 0:	Indels 0:	Gaps 0:
Qy 311	cttttcggtccctccctccgagcctctgtgagccgagatccggccccggaaaccgcgacctga	370		
Db 1	cttttcggtccctctctctccgctccgcttggagccggatccggccccggaaaccgcgacctga	60		
Qy 371	gaagcggtacctctctacgsgttaaggccgttagctctggcggaagagaagagcgccgtctc	430		
Db 61	gaagcggtacctctctacgsgttaaggccgttagctctggcggaagagaagagcgccgtctc	120		
Qy 431	gtcaacagcgccggggagagcccgctgtcttgaggtctggcccgctgcagacctcttcggagc	490		
Db 121	gttcacacagcccggggagagcccgctgtcttgaggtctggcccgctgcagacctcttcggagc	180		
Qy 491	ccagcaatgtaagtcgcggcgagctgcataagaaactccggagccatgtagagatccaggttaa	550		
Db 181	ccagcaatgtaagtcgcggcgagctgcataagaaactccggagccatgtagagatccaggttaa	240		
Qy 551	ggacatctccaggtgtgaatccaaagtcataaaatgaaaaaaacagaccatctctgaaacct	610		
Db 241	ggacatctccaggtgtgaatccaaagtcataaaatgaaaaaaacagaccatctctgaaacct	300		
Qy 611	ctgaaaactgtataacagcgcaagcaaaaaatccaaatgtaagccactctggggaaagtatt	670		
Db 301	ctgaaaactgtataacagcgcaagcaaaaaatccaaatgtaagccactctggggaaagtatt	360		
Qy 671	taccttgacctctcctctctgtacccatatctgaaaaacttcaaaaagacatlaagatctg	730		
Db 361	taccttgacctctcctctctgtacccatatctgaaaaacttcaaaaagacatlaagatctg	420		
Qy 731	ggagggcgagctgtgaagaattctctcagcaaaagatcacgtatcttatcttcaataaagaag	790		
Db 421	ggagggcgagctgtgaagaattctctcagcaaaagatcacgtatcttatcttcaataaagaag	480		
Qy 791	gaagcctaactctgcaaaacctctgtgcgaattctccctgaccaaagtccagaattcgca	850		
Db 481	gaagcctaactctgcaaaacctctgtgcgaattctccctgaccaaagtccagaattcgca	540		
Qy 851	tatctgcgaaagaaacacttaccctcaccacccagcactgctgtaagttaatttaagtcacca	910		
Db 541	tatctgcgaaagaaacacttaccctcaccacccagcactgctgtaagttaatttaagtcacca	600		
Qy 911	gaacaggtgtgtttaagcagagaggaatctatagctgtgaaaaagctacaagagccatgat	970		
Db 601	gaacaggtgtgtttaagcagagaggaatctatagctgtgaaaaagctacaagagccatgat	660		
Qy 971	ttttcttcctcaaatagatataatatacaaaagcctctgtcagctggagagtaaaatcttcat	1030		
Db 661	ttttcttcctcaaatagatataatatacaaaagcctctgtcagctggagagtaaaatcttcat	720		
Qy 1031	atgtagagacatagatataatataatgtaaaagaaaaaagaaggtgtgttttaaccagaaga	1090		
Db 721	atgtagagacatagatataatataatgtaaaagaaaaaagaaggtgtgttttaaccagaaga	780		
Qy 1091	tcaagtaactcagtaagaagatggggcaaaaagcttggtagtgtgtgcataaaaacaaga	1150		
Db 781	tcaagtaactcagtaagaagatggggcaaaaagcttggtagtgtgtgtgcataaaaacaaga	840		
Qy 1151	acaggaagacatcaaaaagcctctgttaaaaggtggagagatataagaccaacttatagcca	1210		
Db 841	acaggaagacatcaaaaagcctctgttaaaaggtggagagatataagaccaactttttatagcca	900		
Qy 1211	ttttatcttaagctgacacataatgctcttataaataatctatctatccagaa	1260		

db	901	TTTTATCTTCAGCTGACCAATATVGGCTTTTATATATATATCTATTCAGAR	950
RESULT	2		
LOCUS	AL581669/c	845 bp	mRNA
DEFINITION	AL581669.1rf_E1011_BCl Homo sapiens cDNA clone CS0DG002ye20.3 prime	EST	16-FEB-2001
ACCESSION	AL581669		
VERSION	AL581669.1	GI:12948899	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	L1.W.B., Gruber,C., Jesse,J. and Polyaer,D.		
JOURNAL	Full-length cDNA libraries and normalization		
COMMENT	Unpublished (2001)		
	Contact: Genoscope		
	Genoscope - Centre National de Sequencage		
	BP 191 91006 EVRY cedex - France		
FEATURES	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.		
source	location/Qualifiers		
	1. .845		

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSDBG002YE20"
/clone_1lb="VRL_F1011_BCl"
/sex="male"
/tissue_type="B cells from Burkitt lymphoma"
/lab_host="DH10B"
/notes="Vector: pCMWSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-stranded cDNA was digested with Not I and cloned
into the Not I and Eco RV sites of the pCMWSPORT 6 vector
Library was constructed by Life technologies. Contact :
Feng Liang Life Technologies, a division of Invitrogen
9800 Medical Center Drive Rockville, Maryland 20850, USA
Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

```

Query Match	Similarity	29.4%	Score 817	DB 103	Length 845
Best Local	Similarity 96.8%	Pred. No. 8,46-123			
Matches 818	Conservative 22	Mismatches 4	Indels 1	Gaps 1	
QY 1683	atccaaacatgagagagacaaatttcctgtataaagagaccaggaactgaaaaaagaac	1742			
DB 845	AMCAACATGAAAGGACGAGATTCTCGTTAAAGACACCCAGAAACGAAAAAAGC	786			
QY 1743	tctctgttttttcagaagccatccccaccccttaaaatgaaattgaagggcttaatgaga	1802			
DB 785	TCTCKTTTATTTCAGAGCCCATGCCCCACCTTCAATGAATTCAGAGGGCTTAATGAGA	726			
QY 1803	aaatgagtaataaagtctcatgtttgtaagtagagcgtgaaatgacataaagacgaatttta	1862			
DB 725	AAATGACTAATAAATCTTCATGTTAAATGACAGTGAAGATGACATAGACAGAAATTTTA	666			
QY 1863	cacagcttaccttaacataaanaaacaagaatgatctctgacattccgaacacacat	1922			
DB 665	CACAGCTACTCTAATATTAANAACAAACAGATTCATTCTTGACATTTCCGAAACACAT	606			
QY 1923	taagtgaataaacttaagaagaactaaaggtagatccataaattgaaataacatagagcat	1982			
DB 605	TAAATGAAATAGCTTTAGAAACCTTAAGGGTATGACTACTATTAATGTGAACATACAGCAT	546			
QY 1983	ctgtaaatgtttctcg-attcagtagacagataatagtgaattctcaacaaagaagaagta	2041			
DB 545	CTGTAACTGTTTCTCAATTTAGTAGACGATAAATAGGATGCTCAACCAAAACGAAAGTCA	486			

```

QY 2042 gatactgctcttccagcaaaaggatctcaagaaaggaagccttcatactact 2101
DB 485 GATACCTGCTTTTCCAGCAAAAGATCTCAAGAAAGAAAGACCTTCATTCATTTTACT 426
QY 2102 catgattctgctcgtatatacaataaagcttcacaagaagcaactcgttccagcaag 2161
DB 425 CATGATTTCTGCTGCTATACATTAACAGTTCACAAAGACACCTAACCTTCAGCAAG 366
QY 2162 gctccatccacatccctccctcgaaggaagcaaggaatgcttgccttcagaataatgtagt 2221
DB 365 GCCCATTCCTMTCCCTCCCTSMSSMCCCAATGATGTSACTTCAAGATATGATAGT 306
QY 2222 ttacctctggttaaaatacacacgaagaaagtgaaataataatagagcaagaaatgaagaa 2281
DB 305 TTMCBCTCTGCTRAATACATCGAAAGAGRRRATATTTAGAGCAAAATGAAAGAA 246
QY 2282 aactcgaagcaaatgctgaatttgataaagaagcaatgaatttatacacaagaagaaac 2341
DB 245 AATCTGGAACCAATGCTGAATTTGATTAAGAACTGAATTTATACACAAAGAAAC 186
QY 2342 agaatttgatctccggtacagctccttactagacttgcttcaagctagtgaaagaa 2401
DB 185 AGAATTTGCTGCTACCGCTACAGCTTACTAGACTTGTTCAGACTAGTGAAGAA 126
QY 2402 tcagaattcttggtctcacaagctacacagaagaagagtgatagtgaaatgtttagat 2461
DB 125 TCAGATTTTGGGTTTACAGCTACACAGAAAGAGTGATATGCAATGTTTAGAT 66
QY 2462 atttggagaaggaataatcagataatcgttcaagcgtttcttcgctccctcaact 2521
DB 65 ATTAGGAAGGAAGAAATTCAGATATCTGTTAACAGCGTTTCTCGTCCCHTCAACT 6
QY 2522 tctac 2526
DB 5 TCTAC 1

```

RESULT 3
AL548911 769 bp mRNA EST 16-FEB-2001
LOCUS AL548911 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D1043Y010 5
DEFINITION prime, mRNA sequence.
ACCESSION AL548911
VERSION AL548911.1 GI:12884384
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 769)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.
Location/Qualifiers
1..769
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1043Y010"
/clone_lib="LTI_NFL006.PL2"
/tissue_type="Placenta"
/note="Vector: pCMVSPORT 6; Site: 1: NotI; 1st strand cDNA was primed with a NotI-Oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :

BASE COUNT 249 a 168 c 170 g 182 t
ORIGIN
http://fulllength.invitrogen.com"
Query Match 27.7%; Score 769; DB 10; Length 769;
Best Local Similarity 100.0%; Pred. No. 5e-115;
Matches 769; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 339 agccggtccgccccgggaagaccgacctgcagagcggttactctcgttagagggc 398
DB 1 ACCCGGATCCGGCCCGGGAACCCGACCTGACGCGGCTACTCTGCTAGAGGC 60
QY 399 gtagcttgaggagaagagagcgccgctcgtcgtcaacagcgccggggaagcgcttct 458
DB 61 GTAGCTGGCGGAGAGAGAGCGCGCGCTCTGTCAACAGCGCGGGAGACCGTCTTT 120
QY 459 cgggctcgccggtgcgaacactctccgagcccaagcagtagtgcggcgactgcga 518
DB 121 CGCGGCTGCGCGGTGGAGACCTTCTCCGACCCAGCATAGTGTGCCGGGAGCTGCCA 180
QY 519 tgaactcggagccatgagatccagagtaagagacattccagggtggaatccaaagta 578
DB 181 TGAAGCTCCGAGCCATGAGGATCCACAGTAAAGACATTTCCAGGGGTGAATCCAAAGTCA 240
QY 579 aaaaagaaaaaacagcacatctctgaaatctcgtgaatcgtataacagccagaanaat 638
DB 241 AAATGTAAGAAAAACAGACCATCTGTAATCTGTGAATCTGTAATACAGCCAGAAAAAT 300
QY 639 ccaatgtgaagccacttggggaagaaagtatttactcctgactactcctgtcacatat 698
DB 301 CCAATGTGAAGCCACTTGGGAAAAAGTATTTACTTACTTACTTCTGTGCACATAT 360
QY 699 ctgaaanaactcaaaagagacataagatcgtggaagcgagltgaagaattctcaga 758
DB 361 CTGAAAAAAGCTCAAAAAGGACATTAAGATCTGGAGGCGGAGTGAAGATTTCTCAGCA 420
QY 759 aagatatacgttatcttattcaataaagaagaagcctaatttgacacaacttgggtc 818
DB 421 AAGATATCAGTTATCTTATTTCAAAATGAAGAAAGAACTTAATTTGACAAACCTTGGGTC 480
QY 819 gaattctcctgtaccaaagtcagaatcgtacatactcagaagaacactcaactcacc 878
DB 481 GAATTTCTCTGTACCAAGTCCAGAAATCTGCATATCTCAGAAACACTTCACTCAGTC 540
QY 879 ccagccatgagtgaagttcaatttaagtcacagacacagtggttgaagcagaagaaat 938
DB 541 CCAGCCATGATGGAAGTTGATTTAACTACACAGACAGCTGTTTAAAGCAGAGAAAT 600
QY 939 tattgttgaagaagctacaaagagacatgatttcttccctcaaatagatatataca 998
DB 601 TATTGATTGTAAGAAAGCTATCAAGGACCATGATTTTATTTCTTCAATATATATTTCA 660
QY 999 atgacctgtcattggggaggaataaattctcatatgataacattagataactcatgaac 1058
DB 661 ATGCCCTTGATGGGAGAGTAAAAATTTCTCATATTTATATACATTAATATCTACATTTGAAC 720
QY 1059 aaaaagaaaaagagtgatattactcaagaatacaagtaactcaagtaag 1107
DB 721 AAAAGAAAAAGAGTGTGATTTACTCAAGAAATCAAGTACTTCAAGTAAG 769

```

RESULT 4
AL573294 784 bp mRNA EST 16-FEB-2001
LOCUS AL573294 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D1043Y010 3
DEFINITION prime, mRNA sequence.
ACCESSION AL573294
VERSION AL573294.1 GI:12932397
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

FEATURES	
COMMENT	Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: sequef@genoscope.cns.fr , Web : www.genoscope.cns.fr Location/Qualifiers
REFERENCE	Mammalia: Eutheria: Primates: Catarrhini, Hominoidea: Homo 1 (bases 1 to 784)
AUTHORS	Li, W.-B., Gruber, C., Jessée, J. and Polayès, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)

```

FEATURES
source
Location/Qualifiers
1. 784
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSOD1043Y010"
/clone_lib="LRI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact: Feng Liang Life Technologies,
a division of Invitrogen 3600 Medical Center Drive
Rockville, Maryland 20850, USA Fax: (1) 301 610 8371
Email: fliang@lifestech.com URL:
http://fulllength.invitrogen.com

```

BASE COUNT	234 a	135 c	126 g	285 t	4 others
ORIGIN					

Query Match	27.68;	Score 768.2;	DB 10;	Length 784;
Best Local Similarity	99.48;	Pred. No. 6.7e-115;		
Matches 778; Conservative	3;	Mismatches 1;	Indels 1;	Gaps 1.

QY	1891	ggaatgcatctcttgacattccgaaacacacatlaagtaagtaaaatgactatgaagaactaag	1950
Db	784	GGATGATTTCTTGACATTTTCCGACACACATTAAGTAATGACTTCGAAAGCACTAAG	725
QY	1951	ggtatgacatactaaatgtaacatacagacatctgcaatgttctcgatttcgaatgaacga	2010
Db	724	GGTGTGATCACTTAAATGTAACTAACTACAGGAGCATCTGTGACATGTGTTCTATTTTCAGTACAA	665
QY	2011	taaatgtagatcttcaacccaacaaacagaagt-cagatactgtgcttltccagcaagatc	2068
Db	664	TAAATGAGATCTCAACCAACACAGAAAGTGCCAGATCTGTGTTTTCCAGCAAAAGATC	605
QY	2070	tcaaggaagaagaccctcatcaatattactcatgattctgtgctgatacaataaaca	2129
Db	604	TCAAGAAAAGACCTTCATTCAATATTTACTCATATTCGTGTCTATTAACAAATAAAC	545
QY	2130	gttacaagaagaccctaactcitttcaaggcaaaagctccatctcatcctccctcgagaac	2188
Db	544	GTTCAACAAAGCACTTAACGTCTTCAGGCAAAAGCCYCCATTCATCTGCTGTGAGAAC	485
QY	2190	ccaatgaatgtgaactccaagaatagtatgtttactctctgtgaaataatcatcgaagaag	2249
Db	484	CCAATGATGTGACTTCAAAGAAATATGATGTGTACCTCTGTGTAATTAATCTCGAAAG	425
QY	2250	tgaataataatattagacgaaatagaaagaanaacttggaaaccaaatgtctgaatttgata	2305
Db	424	TGAAAAATAAATTTAGCAGCAAAATGAAAAAGAAATGTGCAACAAATCTCTGAATTTGATA	365
QY	2310	aaagaacttgaaattatatacacaagaagaacaaacgaatcttgttgcataccggtacagtcct	2365
Db	364	AAAGACTGAAATTTTATACACAAAGAAAGAAACGAATTTGTGATTCAACGGTACAGCTTT	305
QY	2370	tactagaactgtttagagactagtagaagaagaacacgaatatttttgggtttccaaagctaac	2425
Db	304	TACTAGACTGTTTCACTACATAGTAAGAAATACAGATTTTGGGTTTCACAAAGTACA	245
QY	2430	cagaaagaatgtgataatgcaatgttttagatatttgggaagaagaanaattcagataaact	2485
Db	244	CAGAAAGATGGTATATCTCAATGTTTATAGTATTTTGGAAAGAGAAATTCACATTAATC	185

QY	2490	tgaagaagaggttttttcgcgcgccttcaacttacaattacgagcttttaaaatttaa	2549
Db	184	tcttttaacagcgttttttcgcgcgccttcaacttacaattacgagcttttaaaatttaa	125
QY	2550	aaatgcatactcttttcagaagatgaagaatcaattcttgaaattttataaatactga	2609
Db	124	aaatgcatactcttttcagaagatgaagaatcaattcttgaaattttataaatactga	65
QY	2610	tgaagaattcttgagatttttttcagcagcttttttcagagcccaatgtaaatcaana	2669
Db	64	tgaagaattcttgagatttttttcagcagcttttttcagagcccaatgtaaatcaana	5
QY	2670	ata 2672	
Db	4	ata 2	

[illegible]

ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
<i>Enhydra septemvris</i>					
Xenopus laevis					
Humanidae; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
1 (bases 1 to 1019)					
Nih-wgc	http://img.ncbi.nih.gov/nih-wgc				
Unpublished. (1998)					
Contact: Robert Strausberg, Ph.D.					

email: cgaucus@remail.ruhr-uni-gov.de
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraz
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
plate: LNA10737 row: 1 column: 03
High quality sequence stop: 785.

FEATURES
Source

```

/organism="Homo sapiens"
/db.xref="taxon:9606"
/clone IMAGE:485034"
/clone_1lb="NIH_MGC_97"
/lab_host="DH10B"
/notes="Organ: testis; Vector: plasmid (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI; xhoI (gtgaggg
); Oligo-OT primed using primer 5'-TTTTTTTTTTTTTTTTNN-3',
size-selected for average insert size 2.2 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHRI, National
Institutes of Health). Note: this is a NIH_MGC library."

```

Query Match	27.2%	Score 756;	DB 11;	Length 1019;
Best Local Similarity	94.6%;	Pred. No. 6e-113;		
Matches 838; Conservative	0;	Mismatches 40;	Indels 8;	Gaps 5;

Qy 337 ggagccgcatccgcccgggaaaccgacctgcagacgcggtacctctactgctagag 356
|||||
Db 5 GGAGCCGGATCCGGCCCCGGAACCCGACCTGCAGACGGGGTACTCTTACTGCGTAGAG 64

```

QY 397 cgttagctggcggaagagagagagcgccgtctctgtcaaacagccgggggaagcggtgct 456
DB 65 cctgactctggcggaaagagagagcgccgtctctgtcaaacagccgggggaagcggtgct 124
QY 457 ttccggcctggcggagagagagagagagagagagagagagagagagagagagagagag 516
DB 125 ttccggcctggcggagagagagagagagagagagagagagagagagagagagagagag 184
QY 517 catgaactccggagagagagagagagagagagagagagagagagagagagagagagag 576
DB 185 catgaactccggagagagagagagagagagagagagagagagagagagagagagagag 243
QY 577 caaaatgtgaaaaaagagagagagagagagagagagagagagagagagagagagagag 636
DB 244 caaaatgtgaaaaaagagagagagagagagagagagagagagagagagagagagagag 303
QY 637 atccaaatgtgaaaaaagagagagagagagagagagagagagagagagagagagagag 696
DB 304 atccaaatgtgaaaaaagagagagagagagagagagagagagagagagagagagagag 363
QY 697 atctgaaaaactcgaagagagagagagagagagagagagagagagagagagagagag 756
DB 364 atctgaaaaactcgaagagagagagagagagagagagagagagagagagagagagag 423
QY 757 caaagatcatgattatcttattcaataa--gaaggaagctaaatttcgacaaacctg 814
DB 424 caaagatcatgattatcttattcaataa--gaaggaagctaaatttcgacaaacctg 483
QY 815 ggtcgaattctctctgtacaaagctcgaagagagagagagagagagagagagagagag 874
DB 484 ggtcgaattctctctgtacaaagctcgaagagagagagagagagagagagagagagag 543
QY 875 catcccaagcatgagagagagagagagagagagagagagagagagagagagagagagag 934
DB 544 catcccaagcatgagagagagagagagagagagagagagagagagagagagagagagag 603
QY 935 aaattttagtggaaaaagagagagagagagagagagagagagagagagagagagagag 994
DB 604 aaattttagtggaaaaagagagagagagagagagagagagagagagagagagagagag 663
QY 995 tcaaatgacctgtcagggagagagagagagagagagagagagagagagagagagagag 1053
DB 664 tcaaatgacctgtcagggagagagagagagagagagagagagagagagagagagagag 723
QY 1054 tgaacaaagaaagagagagagagagagagagagagagagagagagagagagagag 1112
DB 724 tgaacaaagaaagagagagagagagagagagagagagagagagagagagagagag 783
QY 1113 ggggcaaaagagagagagagagagagagagagagagagagagagagagagagagag 1172
DB 784 ggggcaaaagagagagagagagagagagagagagagagagagagagagagagagag 843
QY 1173 ttgtaaggtggaagagatagagcaacttatagagagagagagagagagagagagag 1218
DB 844 ggcttagcgtgagagagagagagagagagagagagagagagagagagagagagagag 886

```

```

RESULT 6
AL559244 725 bp mRNA EST 16-FEB-2001
LOCUS AL559244.L1.NFL008.TC2 Homo sapiens cDNA clone CSDDJ012YI06.5
DEFINITION prime, mRNA sequence.
ACCESSION AL559244
VERSION AL559244.1 GI:12904555
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 725)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization

```

```

JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segreff@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1..725
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSDDJ012YI06"
/location="L1.NFL008.TC2"
/sex="male"
/tissue="type="T cells from T cell leukemia"
/notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive,
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 224 a 166 c 164 g 171 t
ORIGIN
Query Match 26.0%; Score 723.4; DB 10; Length 725;
Best Local Similarity 99.9%; Pred. No. 1.2e-107;
Matches 724; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 329 cgcgccttgagccggatccggcccggaacccgacctcagaagcgggactctactg 388
DB 1 cgcgccttgagccggatccggcccggaacccgacctcagaagcgggactctactg 60
QY 389 cgtagaagccgtagctgagagagagagagagagagagagagagagagagagagagag 448
DB 61 cgtagaagccgtagctgagagagagagagagagagagagagagagagagagagagag 120
QY 449 gccgtgcttcgcgagctgcgcagagagagagagagagagagagagagagagagagag 508
DB 121 gccgtgcttcgcgagctgcgcagagagagagagagagagagagagagagagagagag 180
QY 509 ggcagctgcagatgaactccgagagagagagagagagagagagagagagagagagagag 568
DB 181 ggcagctgcagatgaactccgagagagagagagagagagagagagagagagagagagag 240
QY 569 atccaagtcacaaatgaaaaaagagagagagagagagagagagagagagagagagagag 628
DB 241 atccaagtcacaaatgaaaaaagagagagagagagagagagagagagagagagagagag 300
QY 629 ccagaaaaatccaatgtaagcactctgggaaagatatttactctgaacttactct 688
DB 301 ccagaaaaatccaatgtaagcactctgggaaagatatttactctgaacttactct 360
QY 689 gtaacatctatgaaaaaactcgaagagagagagagagagagagagagagagagagagag 748
DB 361 gtaacatctatgaaaaaactcgaagagagagagagagagagagagagagagagagagag 420
QY 749 ttcttcagaaagagatcatctatctattcaataaagagagagagagagagagagagagag 808
DB 421 ttcttcagaaagagatcatctatctattcaataaagagagagagagagagagagagagag 480
QY 809 acctgggtgcgaatttctctgtacccaagtcagaatctgatatctgcgaagaacct 868
DB 481 acctgggtgcgaatttctctgtacccaagtcagaatctgatatctgcgaagaacct 540
QY 869 tcaactcatcccgagcatgtagagagagagagagagagagagagagagagagagagagag 928
DB 541 tcaactcatcccgagcatgtagagagagagagagagagagagagagagagagagagagag 600
QY 929 aagaagaaatattagttgaaaaagagagagagagagagagagagagagagagagagagag 988

```

```

Db      601 ACAGACAAATTATTAGTTGAAAAAGCTATCAAGACCATGATTTTATTCCTTCAAAATAGT 660
QY      989 atattcaaatcctctgcacagggagtaaaattcttcattatgagcattgagatc 1048
        |||
Db      661 ATATTATCAAAATGCTTGTGATGGAGATAAATAATCTTCTATTGATGACATTAGATAC 720
QY      1049 tacat 1053
        |||
Db      721 TACAT 725

RESULT 7
LOCUS   AU128881 903 bp mRNA EST 24-OCT-2000
DEFINITION AU128881 NT2RP2 Homo sapiens cDNA clone NT2RP2004396 5', mRNA
sequence.
ACCESSION AU128881
VERSION   AU128881.1 GI:10989235
KEYWORDS EST
SOURCE   human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 903)
AUTHORS  Ota,T., Nishikawa,Y., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
TITLE     HRI human cDNA project
JOURNAL   Unpublished (2000)
COMMENT   Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1533-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3951
Fax: 81-438-52-3952
Email: genomics@hri.co.jp
HRI human cDNA project: 5' - & 3' -end one pass sequencing: Helix
Research Institute: cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute
FEATURES
source
1. 903
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RP2004396"
/clone_id="NT2RP2"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/notes="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
cells after 2-weeks retinoic acid (RA) induction."
BASE COUNT 326 a 165 c 157 g 242 t 13 others
ORIGIN

Query Match 26.0%; Score 721.6; DB 10; Length 903;
Best Local Similarity 92.3%; Pred. No. 2.3e-107;
Matches 820; Conservative 0; Mismatches 56; Indels 12; Gaps 6;

QY      1654 tatcttcagaaagatgtccaggaagatgatatacaacagt-gaagagacagaattcctgt 1712
        |||
Db      1 TATTTCGCAAAAGATGCGACGAAGATGATACAGACGTGGAAGAGCAGAAATTCCTGT 60
QY      1713 ataaagagaccaggaagaaatgaataaaagctcctgttatttcagagccatcccccacc 1772
        |||
Db      61 ATAAAGAGACCCGAGAACTGAAAAAAGCTCCCTGTTTATTTCAGAGCCCATCCCCACC 120
QY      1773 ctccaatgatgtgagggccttaatgagaaatagatataaagtccgttttaagta 1832
        |||
Db      121 CTTAAATGATGAGAGGCTTAATGAGAAAGAGATTAATTAATGTTCCATGTTAACTA 180
QY      1833 cagctgaagatgacataaagacagaattttacacagctacacctacataaaaaaacagg 1892
        |||
Db      181 CAGCTGAAGATGACATMAAGACGAATTTTACACAGCTACCTCTACATAAACAACAGG 240

```

```

QY      1893 aatgactcttgacattccgaaacacacattaaagtaaaaatgacttaagaactaagg 1952
        |||
Db      241 AATGCAATCTTCTTGACATTTCCGACACATTAAAGTGAATAATGACTTAAGAACTAGG 300
QY      1953 taagatcactataaagtacatacagacatcgttacatgtttctgattccagtagata 2012
        |||
Db      301 TAGATCACTATAAAGTAAACATACAGCATCTGACATGTTCTGATTTAGTACAGATA 360
QY      2013 atagatgactacacaaacaaagagatcagatcgtgtctttccacaaagagatcca 2072
        |||
Db      361 ATAGTGATCTCACCMAAACGAAGATCAGTACTGTGCTTTTCCACMAAGATCTCA 420
QY      2073 agaaaagagacttcattcaatatattctcgtatgtctgtctgtgtgtaacataaagatt 2132
        |||
Db      421 AGGAAAGAGACTTCATTCATATATTACTCATGATCTGTGCTGATTAACATTAACAGTT 480
QY      2133 cacaagagacactaactgttccagcaaaagctccatctcactcctcctgaagacc 2192
        |||
Db      481 CACAAGAGACACTTACCTGTTCCAGCNAAGGCTCCATTCATTCCTCTCTGAGAGACCA 540
QY      2193 atgaatgtactcaagaatacgtatgttacccttcgtgtaaaataacatcgaagaatga 2252
        |||
Db      541 ATGAATGTGACTTCAAGAAATATGATATGATTACCTCTGCTGTAATAATCATGAAAGTGA 600
QY      2253 aataatatta-gaacgaataagaaaagaaatcgtgaaccaaactcgtgaattgataaa 2311
        |||
Db      601 AATATATATTAGGAGCAGAAATGAAAGAAATCTGGAACCAATGCTGATTTGATMAA 660
QY      2312 agaactaatatttatacacaagaagaagaactgtgtatgttaccggtacagcttta 2371
        |||
Db      661 ABAATGATATTATTTNCCCCACAGAAACAAACGATTTGT-TTTCACGGGTACAGCTTTA 719
QY      2372 ctgaactgtttcagaactatgaaagaagaacagaattttgggttttcagaagctacaca 2431
        |||
Db      720 CTAAAGCTGTTTCAGACTAGTGAATAAATGCAATTTTGGCTTTCCAAACGCTACCA 779
QY      2432 gaaaa--gagtgatataagcaatgttttagatattgg-----aagagaaatcaga 2484
        |||
Db      780 GAAAAAGAAATTTGATATGCGCATGTTTAAATATATTTGGGGAAAAAAGAAATTTCCNAT 839
QY      2485 taatcgttaacagcgctttt--ctcgtcccttaacttcaacttcaactt 2530
        |||
Db      840 AATCTGTTTACAGCGCTTTTTCCTCCCTCCCTCCACTTCTTCMTT 887

RESULT 8
LOCUS   AL580899 742 bp mRNA EST 16-FEB-2001
DEFINITION AL580899 LTI_NFL008_TC2 Homo sapiens cDNA clone CS0DJ012YL06 3
prime, mRNA sequence.
ACCESSION AL580899
VERSION   AL580899.1 GI:12947367
KEYWORDS EST
SOURCE   human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 742)
AUTHORS  Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE     Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.
FEATURES
source
1. 742
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DJ012YL06"
/clone_id="LTI_NFL008_TC2"
/sex="male"

```


primed. Average insert size 1.229 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC library."

BASE COUNT 275 a 122 c 135 g 180 t
ORIGIN

Query Match 25.6%; Score 712; DB 11; Length 712;
Best Local Similarity 100.0%; Pred. No. 8.4e-106;
Matches 712; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1233 tgcctttataaattatctatctcagaagccctgcagtcacattgatgtagaagccat 1292
DB 1 TGCCCTTTATTAATATTCTATTGAGAGCCCTGCAGTCCATTGATGAGCAAGCCAT 60
OY 1233 ctatgatacaaaagcaaacctcaggttaaaactaagaatccaacagatggcagtaag 1352
DB 61 CTAGATATGCAAAAGCAAACTCAGTTAAACTTAAGATCCAAACAGATGGCATTAAGTATG 120
OY 1353 gtggaacctcaatcaactccaggttgaaagagaagaaaaaagatatgtgaatgt 1412
DB 121 GTGGAACCTCAATTAATCTCAAGTGAAGAGAGAGAGAGAGAGATTTGTGAATGTT 180
OY 1413 gcttcagaagaatagatcagaactcagcttcaagtgaagagacaacagaacttg 1472
DB 181 GCTTGCAGAAATATGAGATGTAGAAACTCAGCTCTTAAGTGACACACAGAAACTTGG 240
OY 1473 caagagataaccagatcagaatgttgatgatattgtatcgaagtgtttgacttg 1532
DB 241 CACAGAGTAACAGATATCAAGTTGTGATGATTTGATCTAGTTAGTTTGTGACTTGG 300
OY 1533 tgggaatgaaagagacacacctaaagaaagaataaatacaggtgttgatccctt 1592
DB 301 TGGATTAATGAAGAGACACCTTAAGAGAGAGAGAGAGATTAAGATGATGATGATGAT 360
OY 1593 ctccgttctcgaagtgctcgaagaaagactgaaacaaagaaagaaagtgatgcaac 1652
DB 361 CTCGCTTTCTGCAAGTGTCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
OY 1653 aatattcagaagaatgtgcagaagatgatatacaacagtgagaagagcaaatctcgt 1712
DB 421 AATATTTCTCAGAAAGATTCGCCAGAGAGATGATACAACTGAGAGAGAGAGAGATTTCTCTGT 480
OY 1713 ataagagagacaggaactgaaaaaagctcgtttatcttcagaagcccatcccccac 1772
DB 481 AATAAGAGACCCGAGAGAACTGAAAAAAGCTCTGTTTATTTTCAGAGCCCATCCGCCAC 540
OY 1773 ctcaaatgaattgagagagcttaatgagaaaaatgataataatgttcattgtaagta 1832
DB 541 CTTCAAATTAATTTGAGAGGCTTTAATGAGAAAAATGAGTAAATGTTTCCATGTTAAGTA 600
OY 1833 cagctgaagatgacataagacagaattttacacagctacactacataaanaaacaacag 1892
DB 601 CAGCTGAGAGATGACATTAAGAGAGAAATTTTACACAGCTACCTGATTAAGAGAGAGAG 660
OY 1893 aatgcatcttgacattccgagacacacataagtgaaaaatgacttagaga 1944
DB 661 AATGCAATCTTGACATTTCCGAGACACACATTAAGTGAAGAAATGACTTAGAGA 712

RESULT 11
BG496289 821 bp mRNA EST 27-MAR-2001
LOCUS 602538259P1 NIH_MGC_59 Homo sapiens cDNA clone IMAGE:4659446 5',
DEFINITION mRNA sequence.
ACCESSION BG496289
VERSION BG496289.1 GI:13457805
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 821)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
COMMENT Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1CM1455 row: e column: 15
High quality sequence stop: 742.

FEATURES

source

1. 821
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4659446"
/clone_lib="NIH_MGC_59"
/tissue_type="mucoepidermoid carcinoma"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: Lung; Vector: pDNR-LIB (Clontech); Site: 1:
SfiI (ggccgctcgcc); Site 2: SfiI (ggccattagcc);
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-ATTCTAGAGCCGAGGGGCGGACATG-dT(30)BN-3-
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH-MGC
Library."

BASE COUNT 273 a 166 c 176 g 206 t
ORIGIN

Query Match 25.5%; Score 710.2; DB 11; Length 821;
Best Local Similarity 96.8%; Pred. No. 1.6e-105;
Matches 799; Conservative 0; Mismatches 18; Indels 8; Gaps 7;

OY 417 gaggcgccgctctctcgaacagcgccgggaagccgtgcttcggagctgcgcgtgcga 476
DB 1 GAGGCGCGCTCTCTCAACAGCGCGGGGAAGCCCTGCTTTCGCGGCTGCCGAGTCCA 60
OY 477 caattctccgagaccagatgtaagtgccggcgagctgcataagctccggagcatga 536
DB 61 CACTTCTCGGAGCCAGCATGTAGTGGCGGGGAGCTGCATGAAGTCCGGAGCATGA 120
OY 537 ggaaccagataaagacattccagggtggaatccaagtcanaaaatgaaaaaacagac 596
DB 121 GGAATCCAGTAAAGGACATTTCCAGAGGTGGAATCCAATGAAAAATGAAAAACAGAC 180
OY 597 catctctgaaatctctgaaactgataaagcgccgaaaaaattccaatgtaagccattt 656
DB 181 CATCTCTGAAATCTCTGAAACTGATTAACAGGCGCAAAAAATCCAAATGTAAAGCCATT 240
OY 657 ggggaaagatatttacccttgactcctctcgtacacatctatctgaaaaactcaaaag 716
DB 241 GGGGAAAAATATTTTACCTTGACTTACCTTCTGTACCAATATCTGAAAAACTTCAAAAG 300
OY 717 acattgaagatctggagagcgagtgaaagaatttccagcaagaatalcagttactta 776
DB 301 ACATTAAAGATCTGGAGAGGCGAGTGAAGAATTTCTCAGCAAAATATCAGTATCTTA 360
OY 777 ttccaataagaagagactaaatttgacacaaacttggtgtgtaatttctctgtaccaa 836
DB 361 TTTCAAAATTAAGAGAGAGTAAATTTGACAAACCTTTGGGTGCAATTTCTCTGTACCA 420
OY 837 gcccaagaatctcatatctgagaagaacacactcactcatalccagcagatgagaagt 896
DB 421 GTCCAGAAATCTGATATCTGCAAGAAACCACTTCACTTATCCAGCAGCATGATGGAAGTT 480

OY	897	catctaaagctccaggaacagcgcgtgctttaagcaagggaaaattatagctctgcaaaaagcta	956
OY	481	caTTTAAgTCACAGACACAGCTGTGTTAAgCAGGGAAATTTAGTGTGAAAAAGCTA	540
OY	957	tcagagccatgattttatctcttccttaaatagatatttcacaaagcc-tctgcacgggga	1015
Db	541	tcAGAGACCATGATTTTATTCTCTTCACAAATAGATATTATCTCAAACGCTTTGCATGGGGA	600
OY	1016	gtaaaaattcttcctatgataga-cattgatactacaccttgaacaaagaanaaaagagtt	1074
Db	601	GTAAAAATCTTCATATTGATGAGACCTTTAGATACACTGATGAAC-AAAGAAACAAAGAGC	659
OY	1075	gattttactccaagaatccaagctctcttcagtaagaatggggcgaanaagagtctgtatgag	1134
Db	660	gATTTACTCAAGAAATCAAGTACTTCCTCAATAGACATGGGGCAAA-GAGCTGTAGTGG	717
OY	1135	tgacacaaaaaacagaacaggaagactccaanaagccttttgaaggtagaagaatagag	1194
Db	718	tgACACAAAAAAGAAACACGAGACACTCAAAAGCCTTTGT-AAAGTGGCAATATGAG	776
OY	1195	ccaacttatagcactttatcttcaagcttgacaaatgtgcctt	1239
Db	777	CCAA-TTTATAGGCCA-TTTATCTTCACAGCAATATATGCTTT	819

RESULT	LOCUS	DEFINITION
12	AI948485/c	
	AI948485	744 bp mRNA
	wg06d08.x1	EST
	similar to TF:075226	08-MAR-2000
		IMAGE:2470479 3
		WUGSC:H.RG135C18.1 PROTEIN ;, mRNA

ACCESSION	AI948485
VERSION	AI948485.1
KEYWORDS	GI:5740795
EST.	
SOURCE	human.
ORGANISM	Homo sapiens

Chromosomes: 10 pairs of autosomes, 2 sex chromosomes (XX or XY).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 1 (bases 1 to 744)
 NC-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 AUTHORS

TITLE	JOURNAL	COMMENT
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index	Unpublished (1997)	Contact: Robert Strausberg, Ph.D. Email: cranbs-remail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA library preparation: M. Bento Soares, Ph.D.
cDNA library arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Centre
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/MLN at: www.bio.litn1.gov/bdbp/image/image.html
Insert Length: 1484 Std Error: 0.00
Seq. primer: 400P from Gibco
High quality sequence stop: 432.

```

FEATURES
SOURCE
LOCATION/Qualifiers
1..744
/organism="Homo sapiens"
/db_xref="taxon:96069"
/clone_image="24/0479"
/clone_lib="NCI-CGAP_Kid12"
/issue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/note="organ: kidney; Vector: pT73D-pac (pharmacia) with a modified polylinker; Site:1: Not i; Site:2: EcoR1; Plasmid DNA from the normalized library NCI-CGAP_Kids was prepared, and ss circles were made in vitro. Following H purification, this DNA was used as tracer in a subcloning hybridization reaction. The driver was PCR-amplified cDNA from a pool of 5,000 clones made from the same library (clonoids 1325912-1325831, 1471368-1472903 and

```

1492104-1493255)	"	Subtraction by Bento Soares and M.
Fátima Bonaldo.		
223 a	127 c	117 g 277 t
BASE COUNT		
ORIGIN		

Query Match	25.58;	Score 710;	DB 10;	Length 744;	
Best Local Similarity	97.38;	Pred. No. 1.8e-105;			
Matches 722; Conservative	0;	Mismatches 20;	Indels 0;	Gaps 0;	

QY 1961 tataaatgtacacatcacggcattctgtacatgtttctctgatttcagttacagataataatgttga 2020

Db 744 TATCAATGTACAAACAGGGCTCGGTACATGTTTCIGATCTCAGTACAGATCATAGTGA 685

Qy 2021 tctcaaccacaacagaagtcagatactgtgcttcttcagcacaagatcctcaaggaaaag 2080

Db 684 TCTCAACCGAGCAGAGTCAGATACTGTGCTTTTCCAGCAGAAGATCTCAAGAACAG 625

2081 gacattcatcaataattctctcatgatcttgctcgtataacaataaacagttcacaaag 2140

Db 624 GACCTTCATTCGATATTTACTCATGATTCCTGGTCGATACAAATAACAGTTCACAAAG 565

QY 2141 caactaactgttcagcgcaaggtccattccataaccctctccctaggaaccaatlgatgt 2200

Db 564 CACCTACTGTT CAGGCAAGGCTCCATTCATACTCCTCGTAGGAACCCCAAGAAATGT 505

QY 2201 gacttcaagaataatgtgacagcttaccctctgtgtaaaatcacatcgaaagcgaaatcaca 445

Db 504 GACTTCAGATATGATAGTTACCTTCCTGGTAAATACATCGAAAGAGGAAATATA 443

[illegible]

Dd 444 TTAGGACGAATAGAAAAAATACTGGACCCAAATCAGTAATTGTTTTCCTACTAGCAATTC 2380

[illegible]

0001 ++++++ctggtcgagggacaaatcttgaattcacaagactacacagggaggagt 2440
db 384 TTTATTCACAAAGAGGAGGGGCGGGAATTGTATCTCAGCCGCAATCTG.....

[illegible]

On 3441 attatattccaatattttatgaatatttgcgaagaagaaattcaqataatctctgttaacagcg 2500

db 264 GGTATATGCAATGTTTAGATATTTGGGAAGGCAAAATTCAGATAATCTGTTAACACGC 205

2501 tttttctctcccttcaactctacatttcaatgcttcttctagaatttaanaaatgcatac 2560

Db 204 TTTTTCCTGCTCCCTTCACATTTACTGGCTTTTGAATTTAAAAATGCATAC 145

2561 ttctcagaactcgtataaqqatcatattctctggaatttttataaatactgtaggaattctt 2620

Db 144 TTTTCAGAGTCGATAGGATCATTTCTTGAAATTTTATAATATGTATGGAATTTCTT 85

2621 agqatatttttacagccttgtttacagaccgaatgtanaataltaaaaaataaatatttg 2680

Db 84 AGGATTTTTTTTACCACTTGTGTACAGACCAATGTAATATTTAAAAATAAATATTG 25

QY 2681 caatttctacagaattgata 2702

Db 24 AAAAAATTCTAAAAA 3

RESULT 13
BG716928
08-MAY-2001

LOCUS	36716928	/03 bp	miRNA	ES1	00	0001	2002
DEFINITION	602689445f1	NIH_MGC_97	Homo sapiens	CDNA clone	IMAGE:4821970	5'	

ACCESSION mRNA sequence.
BG16928

VERSION	BG16928.1	GI:13930113
KEYWORDS	EST.	

SOURCE human.
ORGANISM Homo sapiens

QY	293	gcagcccaactcgtttgtgcttcttgccctctccctccgaccttgagacccgataccgacc	352
Db	62	GCAGCCCACTCCTTTGTGCTTTGGCCCTTCCTCCCTCCGCGCCCTGGAGCCGGATCCGGCC	121
QY	353	ccggaaaccgcaccgcagacagcgagtcaccctactgagctagaagccglaactgaggaag	412
Db	122	CCGGAAACCCTGCACCTGCAGACGCGCTACCTCTACTCTCCGATGAGAGCCGCTAGCTGGCGGAAG	181
QY	413	gagagagggcgccgcctcgctcaaaagaccggggggaagccgctgcttcgcgcgcgcgcgt	472
Db	182	GAGAGAGCGCGCCGCTCGCTCAACAGCCGGGGGAAGCCGTCCTTTCGCGCTCCCGGT	241
QY	473	gcgacacttcctccggaccgcagctgtaagtgccggcgagctgcacatgaactccgagcc	532
Db	242	GCACACTTTCCTCCGACCACGATGATGGTGGCGGCGACTGCATGAACTCCGGAGCC	301
QY	533	atagaagatccacaagaagaagagcattcccaagggagatcccaagtcacaaatctgaanaac	592
Db	302	ATGAGATCCACAGTAAGAAGGACATTTCCAGCTGAGATTCAGTCACCAATGAAAAAAC	361
QY	593	agaccctctcctaactctctgaanaactgataaacggccagaaataatccaatgttaagcca	652
Db	362	AGACCACTCTCAAAATCTCTAANAACGATTAACAGCCGAGAAAATCCAAATGTAAAGCCA	421
QY	653	ctttggggaagaagatcttacccttgacttaacctctgtacacatacttgaaaaacttcaa	712
Db	422	CTTTGGGGAAGAGTATTTTACCTGTGACTTACCTTCTGTACCATATCTGAAAACTTCAA	481
QY	713	aaggacatlaaagatctcggggggcgagatgtgaaga-attctcagcaagaatatctagttia	771
Db	482	AAGGACATTTAAGATCTGGAGGCGCGAGTGAAGAAATTTCTCAGCAAAAGTATTCAGTTA	541
QY	772	tctatttcaaaatgaagaagagagcctaattgacacaaaccttgggtgcgaatctctcgt	831
Db	542	TCTATTTCNAATTAAGAAGAGCTAAATTTGCACAAACCTTGGGTGCAATTTCTCCTGT	601
QY	832	accaagtccaagaatctgcatlaact-gcagaanaacccttcaacctatccagacatgagt	890
Db	602	ACCAAGTCCAGAAATGTGCATTTACTGCGCAAAACCACTTCACCTCATCCAG-CAAGAG	660
QY	891	gaagtcatttlaagccacgcagacacagtggttt-aagcagagaa-aattatagttga	948
Db	661	GAGTTCATTTAAGTCACACACAGCACTGTTTAAGAGAGGAAACAATATTACGTTG	720
QY	949	aaaagctacaagacacatgattttatctcttcataatgattatla	994
Db	721	AAAAATATCAAGACCACTGATTTTATCTCTTCAATAGTATATTA	766
RESULT 15			
LOCUS	AW960753	701 bp	MRNA
DEFINITION	EST72824	MAGE	residues, MGF Homo sapiens cDNA, mRNA sequence.
ACCESSION	AW960753		
VERSION	AW960753.1	GI:8150437	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
	Hedge, P., Ol, R., Abernathy, K., Dnarp, S., Gaspar, R., Gay, C., Holtz,		
	I, E., Saeed, A. I., Sharov, V., Lee, N. H., Yeaman, T. J. and		
	Quackenbush, J.		
	Assessment of gene expression patterns in a model of colon tumor		
	metastasis using a 19,200 element cDNA microarray		
	unpublished (2000)		
JOURNAL	Contact: John Quackenbush		
COMMENT	The Institute for Genomic Research		
	9712 Medical Center Dr., Rockville, MD 20850, USA		
	Tel: 301 838 3528		
	Fax: 301 838 0208		

Email: johne@ligr.org									
Plate: 150									
Seq primer: Reverse.									
FEATURES	Location/Qualifiers								
source	1. 701								
	/organism="Homo sapiens"								
	/db_xref="taxon:9606"								
	/clone_lib="MAGE resequences, MAGP"								
	/note="Vector: pbluescriptSm"								
BASE COUNT	240 a	140 c	147 g	173 t	1 others				
ORIGIN									
Query Match	24.2%	Score 671.4:	DB 10:	Length 701;					
Best local Similarity	98.8%	Pred. No. 3.1e-99;							
Matches 686;	Conservative 0;	Mismatches 7;	Indels 1;	Gaps 1;					
429	ctctcaaacggcgcggaagacgcgcgccttccttcgagctgcgcgcggctgcaaaccttcctcgg	488							
Db	1 ctctcaaacggcgcggaagacgcgcgccttccttcgagctgcgcgcggctgcaaaccttcctcgg	60							
Oy	489 acccagcagctgtaggtgtccgggggagctgcgcagcagaaactccggagccatgaaagcttccagta	548							
Db	61 ACCCAGCAGCTGTAGGTGTCCGGGGAGCTGCAGCAGCAATCCGGAGCAGTAAGATCCACAGTA	120							
Oy	549 aaggacattccagggctggaatcccaagtcacaatatgaaaaaacagacacatctctgaat	608							
Db	121 AAGGACATTTCCAGGCTGGAGATCCAAAGTCAAAAATGAAAAAACAACACCATCTTGAAAT	180							
Oy	609 ctctgaaacatgataacagagccagaaaaaatccaaatgtaagccactctggggaaaagat	668							
Db	181 CTCTGAAAACTGATTAACAGGCCAGAAAAATCCAAATGTAAAGCCACTTTGGGAAAAAGTAT	240							
Oy	669 tttaacctgaactacacctctctgcatactcgtgaaaaacttcaaaaagagacattaaaggatc	728							
Db	241 TTTACCTTGACTTACCTTCTGTCAACCATATCGAAAACTTCAAAAGACATTAAAGATC	300							
Oy	729 tgggaaggcgagcttgaagaatctctcaagaagaagatatacaattcttattcaataaga	788							
Db	301 TGGGAGGGGAGATGGAAGATTTTCTCAGCAAAAGATATCACTTATTTCAAAATAAG	360							
Oy	789 aggaagctaaattgtgcaacaaccttgggtgcgaattctcctgtaccgaagtcgaagatctg	848							
Db	361 AGGAAGCTAAATTTGCACAAACCTTGGTGCGAATTTCTCTGTACCAAGTCCAGATCTG	420							
Oy	849 catatactggagaagaacaccttgcacatcccaacgaagatgaagaattcaatgaagcac	908							
Db	421 CATATCTGCAGAAACCACTTCACTCACTCCAGCCATGATGGAAGTTGATTTAATGCTAC	480							
Oy	909 cagacacagctgtgttaagcagaaggaataattatagttgaaaaagctatcaagaacatg	968							
Db	481 CACACACAGTGTGTTAAGCAGAGGAAAAATTTATGTTGAAAAAGCTATACAGGACCATG	540							
Oy	969 atttattctctcaaatagatataattatcaaalgcctgtgcat-ggggagtaaaaactctt	1027							
Db	541 ATTATTAATCTTCAAAATAGATATATTTATCAATAGCCTTGTCATGCGGGAGTAAAAATCTT	600							
Oy	1028 catatgatgagcatatgatatcatctcatctgcaacaaagaaaagagcttgattactcaag	1087							
Db	601 CATATGATGATGACATTGAATACTCATATGAAACAAAAAGAAAGAAAGTATTACTCAAG	660							
Oy	1088 aaatcaagtaactcagtaagaagaatgggggcaaaa 1121								
Db	661 AAATCAAGTACTTTAATTAACAGATGGGGGCCAAA 694								

THIS PAGE BLANK (USPTO)